

# IDS WORKING PAPER

Volume **2014** No **437**

## **Using Participatory Process Evaluation to Understand the Dynamics of Change in a Nutrition Education Programme**

Andrea Cornwall  
February 2014

Using Participatory Process Evaluation to Understand the Dynamics of Change in a Nutrition Education Programme

Andrea Cornwall

IDS Working Paper 437

First published by the Institute of Development Studies in February 2014

© Institute of Development Studies 2014

ISSN: 2040-0209 ISBN: 978-1-78118-154-6

A catalogue record for this publication is available from the British Library.

All rights reserved. Reproduction, copy, transmission, or translation of any part of this publication may be made only under the following conditions:

- with the prior permission of the publisher; or
- with a licence from the Copyright Licensing Agency Ltd., 90 Tottenham Court Road, London W1P 9HE, UK,  
or from another national licensing agency; or
- under the terms set out below.

This publication is copyright, but may be reproduced by any method without fee for teaching or nonprofit purposes, but not for resale.

Formal permission is required for all such uses, but normally will be granted immediately. For copying in any other circumstances, or for re-use in other publications, or for translation or adaptation, prior written permission must be obtained from the publisher and a fee may be payable.

Available from:

Central Communications, Institute of Development Studies, Brighton BN1 9RE, UK

Tel: +44 (0) 1273 915637 Fax: +44 (0) 1273 621202

E-mail: [bookshop@ids.ac.uk](mailto:bookshop@ids.ac.uk)

Web: [www.ids.ac.uk/publications](http://www.ids.ac.uk/publications)

IDS is a charitable company limited by guarantee and registered in England (No. 877338)

# Using Participatory Process Evaluation to Understand the Dynamics of Change in a Nutrition Education Program

Andrea Cornwall

## Summary

Set against a backdrop of rising concern about the near hegemony of experimental forms of impact evaluation in which all that counts is that which can be counted, this article considers the contribution that process evaluation can make to understanding how change happens. With a focus not just on what changes as an intended result of an intervention, but on a fuller picture of incidental and unintended outcomes, the article describes the development and application of participatory process evaluation to an educational program aimed at addressing chronic malnutrition in an East African country. It explores the use of participatory methods as a tool for learning and accountability that place the “beneficiary” of development at the centre of enquiry and action, and considers the contribution that these methods can make to effective evaluation.

**Keywords:** participatory impact evaluation; process evaluation; randomised controlled trials; Kenya; nutrition

**Andrea Cornwall** is Professor of Anthropology and Development in the School of Global Studies at the University of Sussex. She has worked for many years in participatory development and is a specialist in participatory methodologies. Her publications include *The Participation Reader* (Zed Books, 2011), *Democratising Engagement* (Demos, 2009) and *Spaces for Change? The Politics of Participation in New Democratic Arenas* (Cornwall and Coelho, eds., Zed Books 2007).

# Contents

	Summary, keywords, author note	2
	Acknowledgements	2
	<b>Introduction</b>	<b>4</b>
<b>1</b>	<b>Participatory Process Evaluation: principles and practice</b>	<b>5</b>
<b>2</b>	<b>Listening to stories of change</b>	<b>7</b>
<b>3</b>	<b>Understanding “Success”: Designing a Methodology for Participatory Process Evaluation</b>	<b>9</b>
<b>4</b>	<b>Apocryphal tales and lessons learnt</b>	<b>13</b>
<b>5</b>	<b>Conclusion</b>	<b>20</b>
	<b>References</b>	<b>21</b>

## Acknowledgement

I'd like to thank Irene Guijt, Lawrence Haddad and Eduardo Masset for their very helpful comments on an earlier version of this paper. For their creativity, care and company on what remains the most interesting evaluation experience I've ever had, I'd like to thank Tilly Sellers, John Gashigi and Charity Kibutha. I'd like to express my gratitude to Kirsten Havemann for having commissioned the study on which this paper is based, and for the inspiration I drew from her refreshingly imaginative approach to evaluation.

# Introduction

Rising popularity amongst those who fund international development of the so-called 'gold standard' experimental approach to evaluation has generated growing concern amongst development practitioners. "Rigour" has now become a euphemism for "quantitative". Increasingly, all that counts is what can be counted. Even the most progressive funders now join the clamour for measurable results. Much is at stake. There are evident dangers of over-reliance on a methodology that fails to provide adequately sensitive insights into the processes through which positive change occurs, and at worst is taken as 'proof' of ultimately spurious causalities (Jones et al. 2009). Concern about the implications has manifested in statements, such as that of the European Evaluation Society on 'the importance of a methodologically diverse approach to impact evaluation – specifically with respect to development aid and development interventions' (EES 2007). It has produced initiatives like the Big Push Forward ([www.bigpushforward.net](http://www.bigpushforward.net)), hosts of an April 2013 conference at which prevailing conceptions of 'evidence' were hotly contested. And it has generated a DFID-commissioned study with the explicit remit of exploring alternative modes of knowledge generation on impact beyond the randomised controlled trial (Stern et al. 2012), possibly reflecting some push back from within an institution in which experimental methods were gaining a worryingly hegemonic reach.

The debate on the potentials and limits of experimental approaches to impact evaluation takes us back to some basic questions about what knowledge best serves the purposes of those who fund and support development initiatives. If such interventions are to stand any chance of being sustained, of inspiring others to follow suit and indeed of being successful over the longer term, it is arguably important to understand *how* it is that they work rather than just whether they have achieved those desired results that are amenable to measurement. Bhola distinguishes between three types of impact: 'impact by design, impact by interaction and impact by emergence' (2000:163). Conventional evaluation and impact assessment generally restricts itself to 'impact by design'. Yet there may be significant complications to attribution and causality that arise from 'impact by interaction', described by Bhola as resulting from 'outcomes of an original intervention interacting with other concurrent interventions made by other agents and agencies, and thereby enhancing or inhibiting effects of the original intervention' (2000:163-4). Bhola identifies a third category, 'impact by emergence'. This is especially vulnerable to being completely missed by evaluation and impact assessment approaches that concern themselves only with measuring intended and presumed effects, rather than with tracing the stories of change more inductively. Any methodology that is unable to capture 'impact by emergence' may be missing key factors for success and sustainability (Bhola, 2000; Roche 2000). Bhola hints at the complexity of assessing these impacts, defining 'impact by emergence' as:

...unimagined outcomes emerging from the original intervention through its interactions with other interventions and its interfaces with historical and cultural processes in place but not easily discernable. (2000:161)

In this article, I take up Bhola's concern with 'impact by interaction' and 'impact by emergence'. I draw on the case in which a fuller than usual range of assessment methods were used, including quantitative measurement of change in nutritional status. My focus is one of the more unusual elements of this suite of evaluation approaches: participatory process evaluation. Process Evaluation explores how interventions unfold, on gathering experiential accounts of what went on, what happened that was unexpected and how those involved reacted to those unexpected blights or opportunities. *Participatory Process Evaluation* uses methods that seek to get to grips with the life of an intervention as it is lived and perceived and experienced by different kinds of people, including program or project personnel. It draws on the visualization and engagement methods of Participatory Rural

Appraisal that have been put to such fruitful use in impact assessment (Chambers 1997; Chambers and Mayoux 2005; Guijt 2008a and b).

Participatory process evaluation provides a valuable complement to other approaches to evaluation and impact assessment precisely because it brings into view all three of the kinds of impact identified by Bhola, along with an understanding of how change has come about and how that change might be sustained. In what follows, I draw on primary ethnographic data in the form of notebooks filled with transcripts of interviews and group discussions, pictures of diagrams generated in the process of conducting those interviews and discussions and a field diary. Seeing an intervention through the lens of participatory process evaluation brings into view its internal dynamics and, with this, issues of some sensitivity. I have obscured details of places and people accordingly. What is salient for this analysis is a sufficiently thorough account of the methodology and rigorous enough reflection on its value. I will seek to offer the reader both of these.

## 1 Participatory Process Evaluation: principles and practice

*I tried and failed many times to explain to our donor organisations why processes had an importance beyond the results they achieved. The results-based framework within which we operated existed in the context of complex power relationships... Sometimes we found ourselves talking openly and finding support from among the donors, while at other times we had to conceal our true objectives and ensure that the results-based, logical framework outputs were achieved... We found ourselves adopting a language and a set of tools – technical activity reports, expenditure reports and products – quite distinct from the work we were actually doing...*  
(León 2010, cited in Eyben 2013)<sup>1</sup>

Mainstream international development's concern with ensuring that aid is well spent has led to the adoption of a panoply of tools to produce what Rosalind Eyben (2013:7) calls "results artifacts" and "evidence artifacts". The former consist of 'reporting, tracking and disbursement mechanisms' that include base-line data, performance measurement indicators, reports on results and logical framework analysis. 'Evidence artifacts' include randomized control trials, systematic reviews, cost-effectiveness analysis, business cases and impact evaluation. Eyben observes:

According to 3ie – an organisation created by donor agencies in 2008 to enhance development effectiveness through the promotion of evidence-based policy making - high-quality impact evaluations that measure the net change in outcomes amongst a particular group, or groups, of people that can be attributed to a specific program. They have narrowed the debate from the older OECD Definition of 'impact' which is 'the positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended'. Hence, increasingly impact evaluation (IE) tends to be associated with experimental and quasi-experimental methods because of their 'objectivity'. (Eyben, 2013:20 )

Contestable as this notion of 'objectivity' is for those who would argue that there is no such thing as value-free research and that the frames of reference chosen by the researcher

---

<sup>1</sup> [www.thebrokeronline.eu/Series/Stories-from-Aidland/The-ghost-in-the-aid-machine](http://www.thebrokeronline.eu/Series/Stories-from-Aidland/The-ghost-in-the-aid-machine), 7 December 2010 (accessed 21 April 2013).

privilege certain aspects and preclude others, this kind of narrowing of gaze is an example of the rampant reductionism that has crept into impact assessment in recent years. Only what can be counted, it seems, counts as 'evidence'. Yet measurable net outcomes, however they are counted, tell us very little about *how* change happens. This is well recognised in the writings of the Executive Director Howard White on the use of combined methods (Carvalho and White 1997; White 2008, 2011), including participatory methods and participant observation; as White puts it, "measurement is not evaluation" (2011: 132). Conventional quantitative impact evaluation also provides us with few insights into how change can be *sustained* (White 2011), or indeed, as Michael Woolcock notes, into the temporal dimensions of impact:

Understanding the efficacy of development projects requires not only a plausible counterfactual, but an appropriate match between the shape of impact trajectory over time and the deployment of a corresponding array of research tools capable of empirically discerning such a trajectory. At present, however, the development community knows very little, other than by implicit assumption, about the expected shape of the impact trajectory from *any* given sector or project type, and as such is prone to routinely making attribution errors. Randomisation per se does not solve this problem. (2009:2)

Woolcock directs us to looking more carefully at impact trajectories. He is roundly dismissive of the linearity implicit in prevailing experimental approaches to impact evaluation:

It is only the most ad hoc theorising or wishful thinking (or the overriding imperatives of domestic political cycles and the structure of career paths at development organisations) that could possibly substantiate an assumption that *all* project impacts are linear and monotonic. (2009:3)

Instead, he notes that despite the recognition that impacts may vary in different settings and with different groups of people, surprisingly little attention is paid to temporal variability. He goes on to argue

... efforts to enhance development effectiveness through evidence derived from project evaluation need to move beyond debates pertaining to the 'rigour' of isolated methods, to more concerted attempts to understand mechanisms driving impact trajectories over time, in different places, at different scales, and in accordance with how well they are implemented. (2009:15)

Going back to the OECD definition cited by Eyben, if 'impact' is taken to refer not to countable things but to 'the positive and negative changes produced by a development intervention, directly or indirectly, intended or unintended', what one would hope an impact assessment methodology to be able to do is to get behind the things that have come out of an intervention to understand the nature of those changes. As well as reflecting on how intended and direct changes actually play out on the ground, such a methodology would be most useful if it could also identify changes that are *indirect* and *unintended*, as these offer often important insights into issues like sustainability, replicability and the potential for diffusion. Taking up Woolcock's argument, if such a methodology could also enable us to understand in some depth the patterns of change over time that are giving rise to sought-after or unexpected outcomes, it might just provide the means for addressing the challenge of making impact evaluation properly rigorous. These are dimensions of impact assessment that the current focus on experimental methods in impact evaluation does not and indeed *cannot* adequately address (Ravallion 2009; Cartwright and Munro 2010).

What makes RCTs useful is that they are able to address the highly targeted question of 'does this intervention work in this place, with these people, at this time'. From this

inferences might be drawn about whether an intervention would work elsewhere, with other people, at another time. But the RCT itself does not provide researchers with sufficiently rigorous contextual evidence to permit them to answer questions of substantial importance. Stern et al. (2012: 8) comment:

Jonas et al (2009) observes that RCTs can answer ‘does it work?’ questions but not ‘how could it be improved?’. Cartwright and Munro (2010: 265) note that experiments can answer the question ‘did it work here?’ but not ‘will it work for us (elsewhere)?’

Participatory process evaluation takes as its starting point an interest in precisely these questions, and is explicitly concerned with mapping ‘impact trajectories’ over time in relation to different dimensions of an intervention and its outcomes. It is built on the recognition that perceptions of ‘progress’ and ‘success’ vary over time and between different stakeholders. It is premised on the understanding that at the outset, expectations of what the intervention could or should achieve may be quite different from how things end up once implementation is underway, and that along the way, views may change as a result of the process, and stakeholders may end up in quite a different place than they originally envisaged – and with a different understanding of what ‘success’ means. Rather than simply look at ‘results’, participatory process evaluation seeks to contextualise the process of intervention in relationships and institutions, looking at shifting institutional frameworks and relational dynamics. Temporal narratives generated through the use of visualisation methods aimed at facilitating story telling provide opportunities to explore critical moments: crises, breakthroughs, conflict, gradual and sudden change. Analysis of the past and of different experiences of transition to the present open up opportunities to reflect collectively on desired changes for the future, working from the perspective of different actors in the process to facilitate a dialogue on futures possible.

By seeking to understand how change happens from the perspective of a diversity of people involved in a particular situation, participatory process evaluation can generate a rich picture of the preconditions for positive change, as well as a deeper understanding of obstacles or barriers. By looking for change not only in the places towards which deliberate efforts to bring about change have been directed, but in the positive changes people report experiencing – that is, rather than measuring progress against a baseline according to a stock of indicators, looking more inductively at understanding what’s changed, how and why – participatory process evaluation brings into view unexpected and unpredicted changes, as well as providing a basis for reflection on why and how what was achieved happened or did not happen. It provides a way of getting to grips with all three of Bhola’s dimensions of impact assessment and, as such, a valuable complement and corrective to quantitative methods of impact evaluation in general and RCTs in particular.

## 2 Listening to stories of change

Two years into an integrated government nutrition program that sought to address pervasive malnutrition and worrying levels of stunting in a rural community in central Kenya, the program’s funders – a progressive Nordic donor – decided it was time to take stock. A baseline survey had been carried out. Social marketing had carried the results of the survey into communities, and participatory methods had been used to engage community members in analyzing the situation that had led to such poor nutritional outcomes for their children. A cluster of interventions had been developed that included educating families on spotting the signs of malnutrition and a host of preventive actions such as hand washing, dish racks and latrines, and making the leafy green vegetables that had come to be regarded as the food of the poorest into a desirable food for all. Outreach from a feeding centre and adult education in organic farming and income-generating skills was complemented with community-based



participation activities and nutrition and hygiene education work with schools. Small business, farming and health projects had been established as part of a holistic, multi-sector approach. Methods used for training and teaching adults and children ranged from more didactic approaches, to the use of Participatory Educational Theatre, Child-to-Child and Participatory Rural Appraisal (PRA) to visualize the problem, examine the causes, publicise the consequences and create a space for public engagement in finding solutions to malnutrition and stunting.

By the time of the process evaluation, two years into the initiative, monitoring of nutritional status was already showing the kind of results that would materialize in the impact assessment that would be carried out five years after the program began and would demonstrate the success of the program in bringing about significant changes in nutritional status amongst children aged 12-60 months. This study, using experimental methods, showed significant reductions in the levels of underweight (mean Z-score:  $-1.66$  v.  $-1.37$  ( $P < 0.02$ ); % with Z-score  $< -2$ :  $42.9\%$  v.  $31.4\%$  ( $P < 0.035$ )) and stunting (mean Z-score:  $-2.05$  v.  $-1.59$  ( $P < 0.05$ ); % with Z-score  $< -2$ :  $52.7\%$  v.  $39.7\%$  ( $P < 0.02$ )), compared to little change in communities without intervention (mean Z-score:  $-1.63$  v.  $-1.50$  (NS); % with Z-score  $< -2$ :  $36.6\%$  v.  $34.5\%$  (NS)) and stunting (mean Z-score:  $-2.0$  v.  $-1.99$  (NS); % with Z-score  $< -2$ :  $44.3\%$  and  $47.4\%$  (NS)) at baseline and after three years (Havermann et al. 2012).

It didn't take a sophisticated anthropometric study, however, to demonstrate some of the more tangible results of the program: any visitor to the area could observe gardens with leafy green vegetables, dish racks and latrines. If that visitor was an outsider and was taken to a school, they would be treated to displays of singing featuring hand-washing and other health promotion activities that could reduce the incidence of diarrheal disease. The old feeding centre used in the past to rehabilitate starving children had turned into a lively hub of activity supporting organic farming, income-generating projects and a network of community participation resource persons. Things were clearly working.

But how did the program achieve those results? How did it actually work – what made it tick, what made it successful? And what was there to be learnt about the way in which things were done that might be useful for other places, other problems? What explained the success that the program had been able to have in enlisting community participation and engagement from local government? How sustainable was the injection of intensive effort that had accompanied the rolling out of the program in this district? And what wider lessons might be learnt from taking a closer look at some of the ways of working that had been developed by the program including their use of unconventional approaches such as theatre and participatory visualization techniques? For this, the donor felt that what was needed was a *process* evaluation, to accompany a more conventional program review. And since participatory methodologies had been so powerful a part of the learning journey in the implementation of the program, the decision was made by the principal adviser to the program to commission a participatory process evaluation. A team was brought together consisting of two Kenyan participation experts, Charity Kibutha and John Gashigi, and two British consultants with experience in the use of participatory approaches, Tilly Sellers and myself.

Over the period of about ten days, we worked intensively together to develop and apply a methodology for participatory process evaluation. In doing so, we drew on our accumulated collective experience of using participatory methods in evaluations of various kinds; in the process, we invented a sequence of activities that adapted existing methods, putting them to use to explore the questions with which we were interested, in an iterative process that took shape as we worked. Unaware of Bhola's work, which provides such a very useful framing for this reflection on method, our interest was in getting to grips with the dynamics of change in all its complexity, and especially in understanding the relational dimensions of how these

changes had been brought about. In what follows, I narrate the process of developing and applying the methodology. I reflect on what we learnt from this experience, exploring the broader implications of the use of participatory process evaluation in impact assessment to 'reach the parts' that experimental methodologies are largely unable to reach.

### 3 Understanding 'Success': Designing a Methodology for Participatory Process Evaluation

It was not long after arriving in the area that we were to show the program management team quite how different our approach to evaluation was going to be. After a briefing by the program leader, which outlined the major activities of the program and gave us some of the background to the nutrition situation in the area and the kinds of program interventions that had been pursued to address it, we were informed as to the fieldsites that had been selected for us to visit. Half an hour under a tree in the yard with one of the extension workers was all it took to elicit a comprehensive matrix ranking of sites, using a series of criteria for success generated by the extension worker, on which three sites we had been offered as a "range" of examples appeared clustered at the very top of the list and one at the very bottom. A quick cross-check with a variety of differently positioned program staff yielded a broad confirmation of his assessment, along with some surprise at how we'd come about this information hours into our visit.

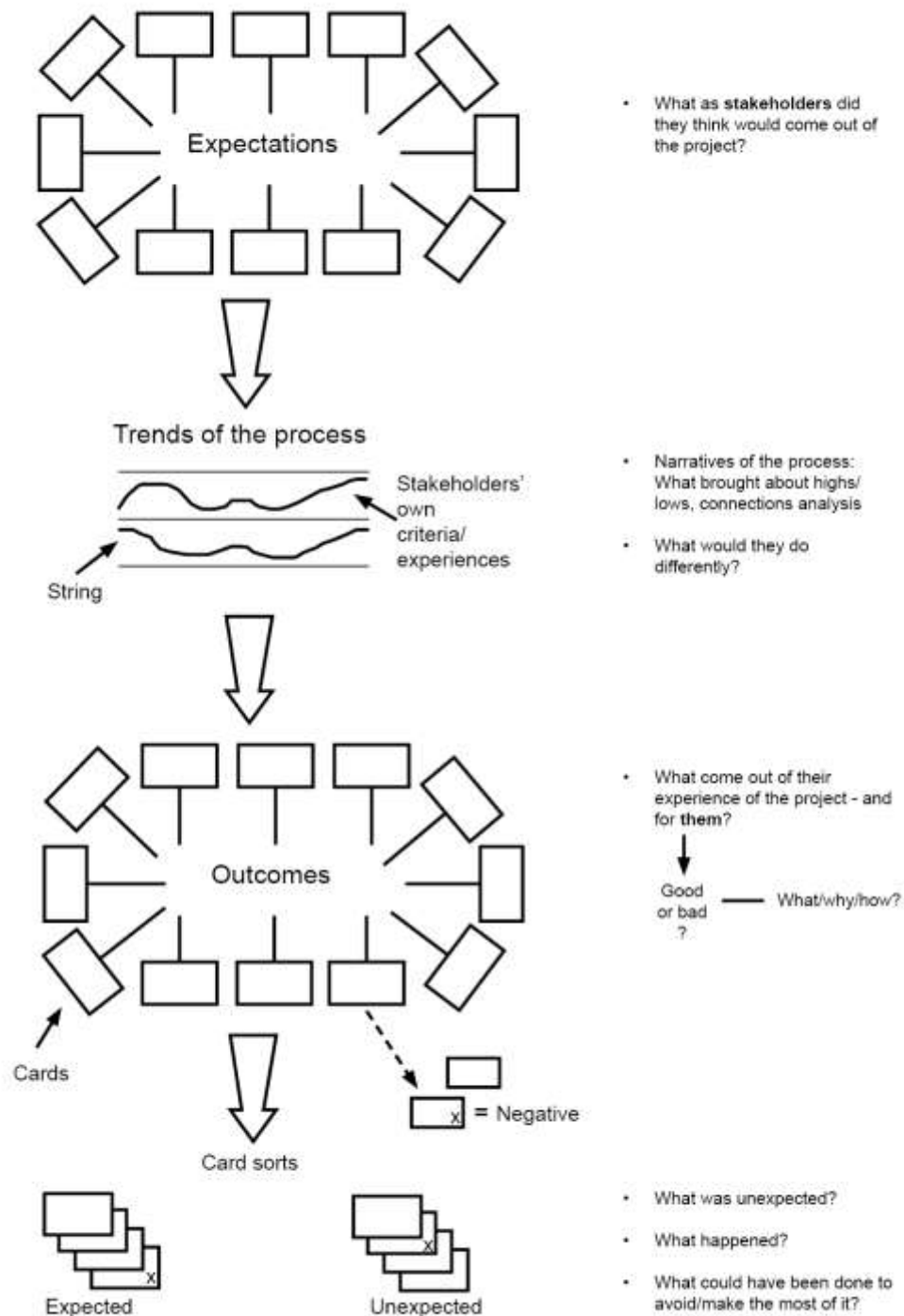
That we were being offered this pick of locations was, of course, to be expected. Evaluators are often shown the showcase, and having a basket case thrown in there for good measure – a case where nothing more might have been possible – allows success stories to shine more brilliantly. Even though there was no doubt in anyone's minds that this was an exceptionally successful program, our team had been appointed by the donor responsible for funding; it was quite understandable that those responsible for the program were taking no chances. It was therefore with a rather bewildered look of surprise that our request to the program manager was greeted to visit a named list of sites, chosen at random from various parts of the ranked list, and not the ones we were originally due to visit.

What we did next was not to go to the program sites, but spend some more time in the headquarters. We were interested in the perspectives of a variety of people involved with implementation, from the managers to those involved in everyday activities on the ground. And what we sought to understand was not only what the project had achieved, but also what people had *thought* it might do at the outset and what they *hadn't expected* to happen, and what had *surprised* them (cf. Guijt 2008a). Asking these kinds of questions gave us vital clues into aspects of how the program was working that were hidden, intangible and yet absolutely crucial. It also allowed us to understand the institutional dynamics of the program in ways that simply could not have surfaced if all that we were interested in were results in terms of nutritional status, with – again – substantial implications in terms of the sustainability of its successes. Rather than flatten the polyphony of perspectives on program into a set of outcome indicators that could be quantified, we actively sought out people's versions of what had happened and of how the changes that were significant to them had come about. These 'stories of change' offered us a more robust, rigorous and reliable source of evidence than the single stories that conventional quantitative impact evaluation tends to generate. They allowed us to explore a richer picture of what "success" might mean to different kinds of people, one that would provide opportunities along the way for reflection that could spur further action.

Our methodology consisted of three basic parts. The first was to carry out a stakeholder analysis that allowed us to get a picture of who was involved in the program. We were interested in hearing the perspectives not just of program “beneficiaries”, but also of others - everyone who had a role in the design, management and implementation of activities, from officials in the capital to teachers in local schools. This involved a series of interviews in which we used coloured cards to generate a visual representation of everyone involved in the conception, design and delivery of the program, which we also used to do some analysis on the respondents’ perceptions of the relative importance and relationships between these stakeholders, using an adapted Venn Diagram method for institutional and relationship mapping. On the basis of this, we made an initial selection of people and groups to interview.

The next step involved a very simple sequence of exercises that we devised to capture expectations and experiences over the life of the program, using a packet of coloured cards, a large piece of paper and pieces of string. It began with an open-ended question about what the person or group had expected to come out of the program. Each of the points that came out of this were written by one of the facilitation team on a card, one point per card, and extensive prompting was used to elicit as many expectations as possible. Once the expectations had stopped flowing, the next step of the process was to look at what was on the cards and cluster them into categories. Each of these categories then formed the basis for the next step of the analysis, which was to look at fluctuations over time. This was done by using the pieces of string to form a graphical representation, with the two-year time span on the x axis and points between two horizontal lines representing the highest and lowest points for every criterion on the y axis. What we were interested in was the trajectories of those lines - the highs and lows, the steady improvement or decline and where things had stayed the same, whether not budging or being sustained over time. We encouraged people to use this diagram as a way of telling the story of the program, probing for more detail where a positive or negative shift was reported. Fascinating details emerged which we then probed further on to reveal dynamics, strategies and tactics, and successes and crises that might otherwise not have come into view.

**Figure 1: Methodological Design**



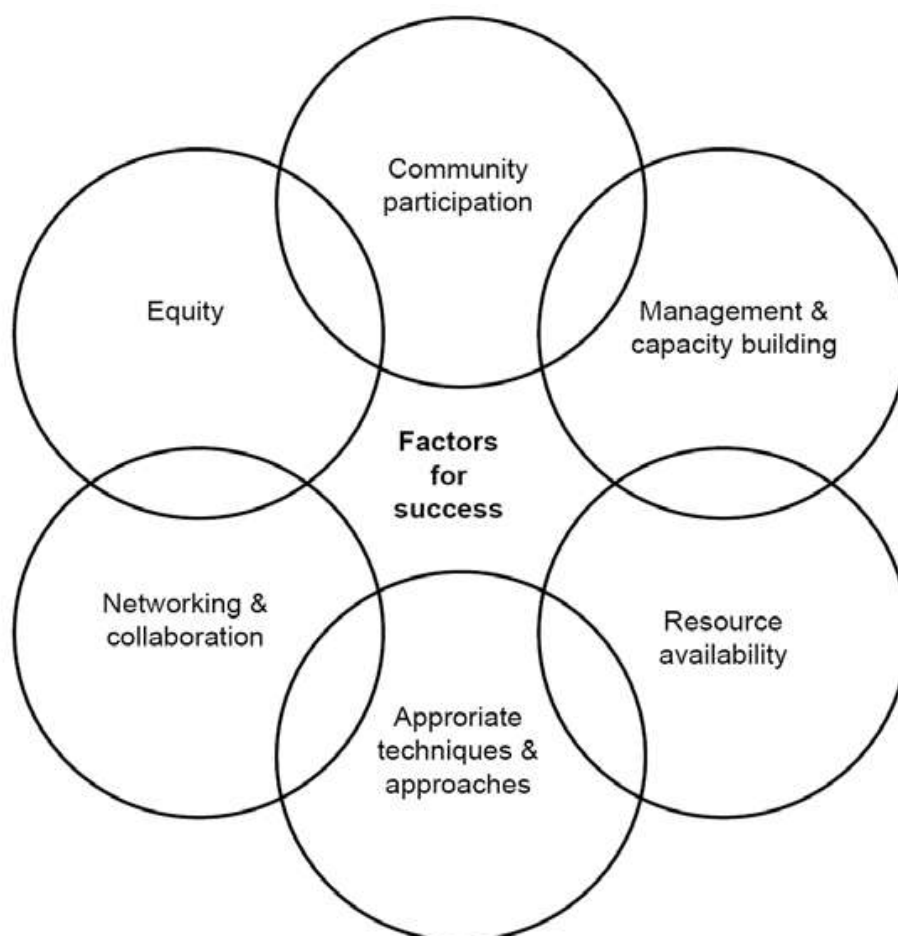
The third step was to use this data as a springboard to analyse and reflect on what had come out of their experience of the project. We did this by probing for positive and negative outcomes. These were written onto cards. We asked people to sort them into two piles: those that had been expected, and those that were unexpected. We then spent some time reflecting on what emerged from this, focusing in particular on what could have been done to avoid or make more of unexpected outcomes and on the gaps, where they emerged,

between people's expectations and what had actually happened, building on the stories that had emerged from the time trends diagrams. We kept people focused on their own experience, rather than engaging in a more generalized assessment of the program, and sought a reflective dialogue about what went well and what might have been done differently, prompting people to think about what might be done differently if the program were to be run again in another place.

We complemented this sequence of methods with some basic finding out using PRA methods in villages in the area, with children and adults. We used social mapping to explore who was involved in the program in the village, and to map institutions; we used this to open up discussion on how the program was working at the village level, who was participating in it and who was perceived by villagers to benefit from it. Venn diagrams representing the period prior to the project and the present were used as entry points for discussions about processes of institutional change, charting the response of existing institutions to the activities of the program and looking at dynamics between institutions over the period of the program. We used network diagramming to explore connections between individuals and institutions that had some relationship with the program, analysing the degree of functionality of these networks in terms of communication and support and points of tension, competition and dysfunction. Time trends of community participation were used to generate discussion on how people have been involved, on what community participation means to people, and to explore people's experiences of barriers, obstacles, conflicts and successes since the program began. Group discussion generated a series of common criteria representing what villagers felt to be vital to securing improvements in nutrition, which were then applied to ten villages: each criterion was scored and an overall score given, from which a discussion was then facilitated on factors for success. This was then used as the basis from which to extend the analysis beyond analysing the change that had happened to sustainability and replication. From this we generated a diagrammatic representation visualising factors for success that was subsequently used as an entry point for further discussion, and for developing process indicators against which the program could be monitored and evaluated.

Rather than simply gathering and analysing the data ourselves, we enlisted the participation of different stakeholders in the process of sense-making and shared with them the emerging picture. This included the external review team, with whom our paths crossed a couple of times. The first time, we exchanged stories of what we'd been doing: theirs of being driven to meetings at chosen sites to meet "the community" and sitting politely through hours of singing and speeches, ours of sitting in people's homes or talking to them in their fields, hanging out with groups of children and government or project staff. Comparing notes, it became evident that we were in a completely different position in terms of understanding the nuts and bolts of how the program worked. The second encounter was close to the end of the review period, when they came to find out what we'd learnt. With the simple charts and diagrams we'd produced, retelling the story was swift – and safe for those who shared their confidences, the PRA diagrams providing a cloak of anonymity and our synthesis of headlines offering the kind of discretion that allowed us to surface emerging challenges without revealing our sources.

**Figure 2: Factors for Success**



Our way of working left a particular impression with the district officials we spent time with – partly, I suspect, because they were not often asked their views by the stream of outside visitors to the program, let alone invited to talk about their experiences and feelings. On our last night in the community, we were invited to have a drink with some of these officers. “We’ve never had visitors coming here who knew so much,” one said to us. Another confided that it’s easy enough to direct the usual kind of visitor towards the story that the program team wanted them to hear. Development tourists, after all, stay such a short time: “they’re in such a rush, they go to a village and say they must leave for Nairobi by 3 and they [the program staff] take them to all the best villages.”

## 4 Apocryphal tales and lessons learnt

One of the most powerful lessons that the program learnt came from a very unexpected reaction to something that was so utterly conventional, it would not even have raised the slightest suspicion that the events would unfold as they did: a baseline survey. A team of enumerators had set out to gather data from a random sample of households, which included basic demographic and anthropometric data such as height for weight and upper arm circumference measurements. At the same time as this survey was being conducted, a rumour was sweeping the area about a cult of devil-worshippers seeking children to sacrifice. Families greeted the enumerators with hostility. People in the communities likened the measurement kits developed for ease of use to measuring up their children for coffins. The survey proved difficult to administer. In one place, the team were chased with stones.

To get things off the ground again, the program needed the intercession and the authority of the area's chiefs to call their people and explain what the program was all about and what it was going to do for the area.

What was so striking about the stories of this initial process of stumbling and having to rethink was that it simply had not occurred to the researchers that entering communities to measure small children might be perceived as problematic. They had explained the purpose of the research and sought consent. They had had their interviewing protocols. They had official documentation. None of this mattered. The incommensurability of knowledge systems, the lack of trust of outsiders and the failure to adequately countenance what was needed to establish the basis for informed consent provoke questions participatory researchers have raised for decades about the nature of 'extractive' research and the politics of 'expert' knowledge (Gaventa 1993, Tandon 1996).

The apocryphal tale of the baseline survey offers other lessons. Conventional research is rarely transparent in any way. It may not be at all evident to anyone why certain questions are being asked, why certain people have been chosen to answer them, what's going to happen to their answers and what the knock-on effects of taking part might be. It is quite reasonable, under these circumstances, to do as Wolf (1987) reports in his classic article "Lying Informants": pull the wool over the enumerator's eyes. Participatory Rural Appraisal methods offer a degree of transparency and engagement that the standard survey method lacks. This is not to say that the versions that emerge from diagramming methods are not also inflected with a reading of what outsiders might wish to hear (Cornwall 1996). Yet the very public nature of these processes and the kind of explanations of purpose that take place, as well as the opportunities for engagement in analysis that they present, does change the dynamic; this can produce outcomes that are more trustworthy precisely because of the kind of cross-checking and deliberation that happens as part of the process (Chambers 1997; Pretty 2005).

Gathering together groups of people to discuss problems and opportunities provided the program with an entry point that helped to re-establish trust with communities in the area after the baseline. Using social marketing techniques to communicate the findings from the survey provided a way to explain to people why the survey was needed, and to make sure that they understood the implications of the results. One community member shared with us, "when X [the feeding centre people] came, many didn't like it because they didn't know what was going to happen". The PRA activities conducted by the program provided an entry point for communication. Activities like mapping exercises gave people new insights into the place where they lived. Sitting together with the community, program staff facilitated, documented, synthesised to get to the main problems. These problems were then sorted into categories, focusing on where the opportunities were to do something. It took a while for it to sink in that this broader problem analysis and work on systemic solutions could address the problems the community were experiencing, as one villager shared with us:

When we started with the activities, I thought what has this got to do with malnutrition – there is still no food. But later I came to realise that there are lots of causes of malnutrition like disease, water and not eating the right foods.

Recollecting his experience of implementing the PRA activities, a program worker talked of how people came back from these kinds of activities talking about how they 'saw' malnutrition, citing one of the villagers who had said, 'when you presented the data, it didn't make sense – now it makes sense.' Comparing crops and stunting in children helped get the message across; the very possibility that malnutrition could affect the children's brains and make them unable to study was compelling a reason to care enough to get involved in the program in a country where education was so highly valued. 'We learnt many things we didn't know', we were told in one village. In another, a village resource person told us of the

surprises this had brought, 'we didn't expect we would learn this much or there was so much that we didn't know.' He went on, "the spread of information [now] is such that it penetrates every part, from one to the next person, there is no way we can go back now, just forward.' As another program worker recalled, "in some cases, we felt almost left out... it was so exciting to see them take over. The literacy level is not the problem, they're able to understand and use the information themselves."

What came out of the program's participatory community-level work and from the social marketing was striking: there was very little awareness of quite how bad the situation was. The shock with which the data was received spurred further reflection amongst those who lived in these communities on what had gone so very wrong for so many children to be suffering in this way, and served as a powerful prompt for actions that could address a problem that was now shared. By using participatory diagnostic methods to uncover problems and explore solutions, then using participatory educational methods – Participatory Educational Theatre, Child-to-Child, and PRA – to engage people in exploring causes, consequences and solutions that they themselves could become engaged with, the program was able to effectively engage participatory research as pedagogy. Community-led growth monitoring and a community evaluation process gave community members access to learning other tools that they could put to work in assessing progress.

Reflecting on the 'impact trajectory' of the program from the distinctive vantage points of different stakeholders offered us fascinating insights. Discussions with key co-ordination personnel in government (to whom the management of the program was shortly due to devolve once the intermediary implementing agency withdrew), for example, yielded six factors that were most significant in shaping their expectations and experience of the program: sense of ownership of the program; how the community viewed them; level of confidence they felt they had in the work; level of demoralization associated with the program; level of their own participation and engagement with it; level of threat that they felt the program posed to them. These emerged from meta-analysis of the cards generated from the first step of the time-line exercise, and provided the basis on which the lines were subsequently constructed. As can be seen from this list, this is not at all the conventional stuff of an evaluation that simply looks at what has been achieved in terms of outcomes 'on the ground'. Yet in terms of the continuity of the program, the story they had to tell was a very interesting one.

At the beginning, they said, they felt very little sense of ownership of the program. It took almost a year before they began to feel involved, and then things got stuck: two years down the line, things had not really moved much further than this, and there were issues in terms of access and control that were clearly still fairly sticky. But there had been a dramatic change in how they felt the community viewed them, as government: their graph shot up quickly from the onset of the program, and continued at a very high level. This had a knock-on effect on their morale. They felt more confident in their work, and their level of demoralization, which was very high at the outset, went right down, in incremental bounds. Riding on all of this was their own participation in the program, which rose exponentially and plateaued at a high level by the second year. And lastly, in inverse proportion to this, was the level of threat they felt, which began at the highest of levels and had all but disappeared just over a year into the program.

Their colleagues at the divisional level had an equally positive story to tell, although one that was tinged with similar frustrations to those experienced by community volunteers and resource people. Their themes were: negative or positive attitude to the project; morale; resources available to them; standard of living (theirs). At the outset, they were pretty hostile to the program. It didn't seem to be offering them much in the way of resources, especially for fuel for their vehicles to enable them to go out and about in the communities. They were concerned about the impact of rising expectations of them. This negativity fell steadily over



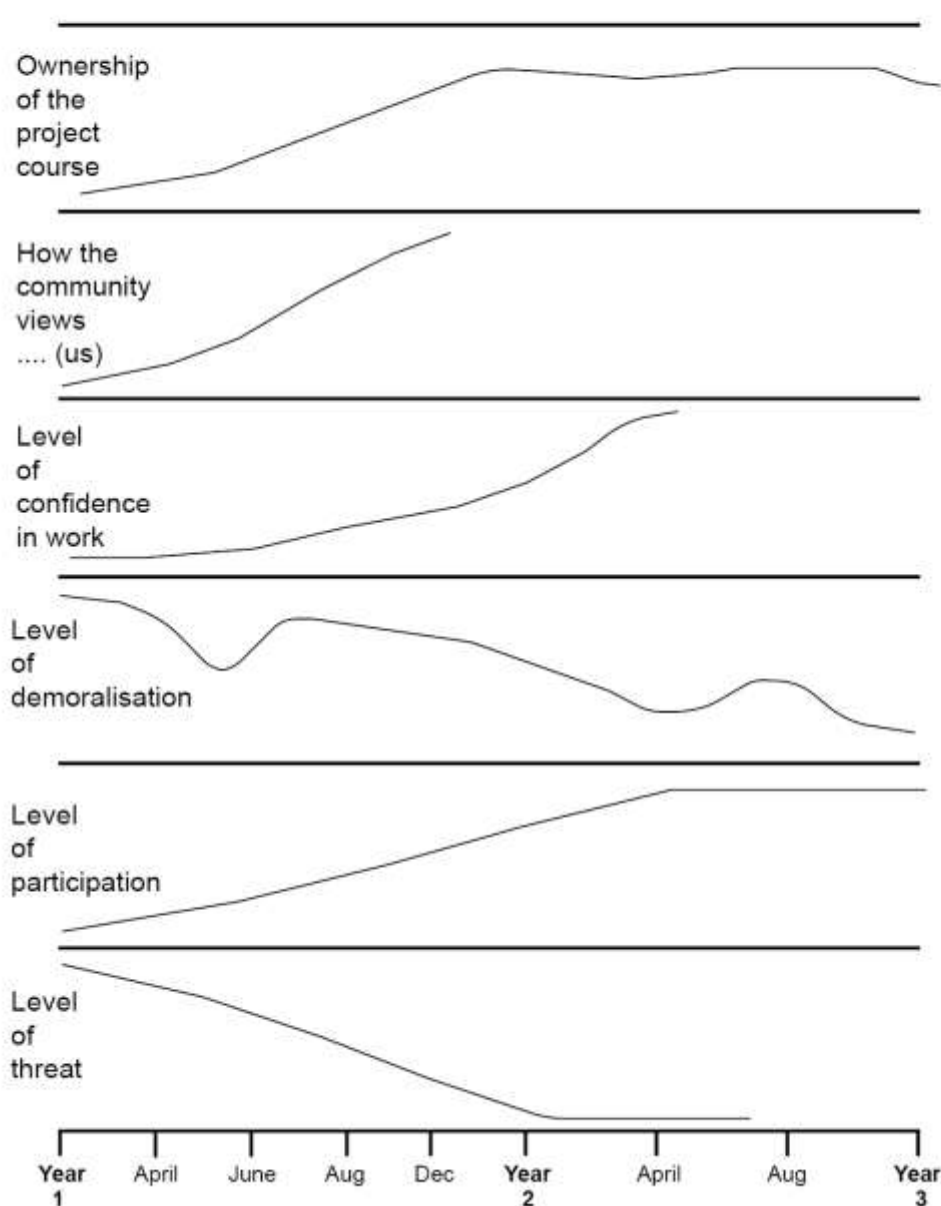
the course of the program, changing direction in the last few months to a general positive sense of what the program offered them. Yet while at the start, their morale had risen as their negativity had dipped, this hadn't been sustained: the same ownership and control issues arose and the same resource issues that were raised by those working in the community hit them in the middle of it all. The resources graph registered a small rise and then a fall, representing allowances they were initially paid and the dissipating hope that further resources for their work would be available. Lastly, their standard of living registered an incremental, but small rise – far less than they had initially expected.

**Figure 3: Time Trends: Key governmental co-ordination personnel**

**Time trends:**

Key co-ordination personnel at government level.

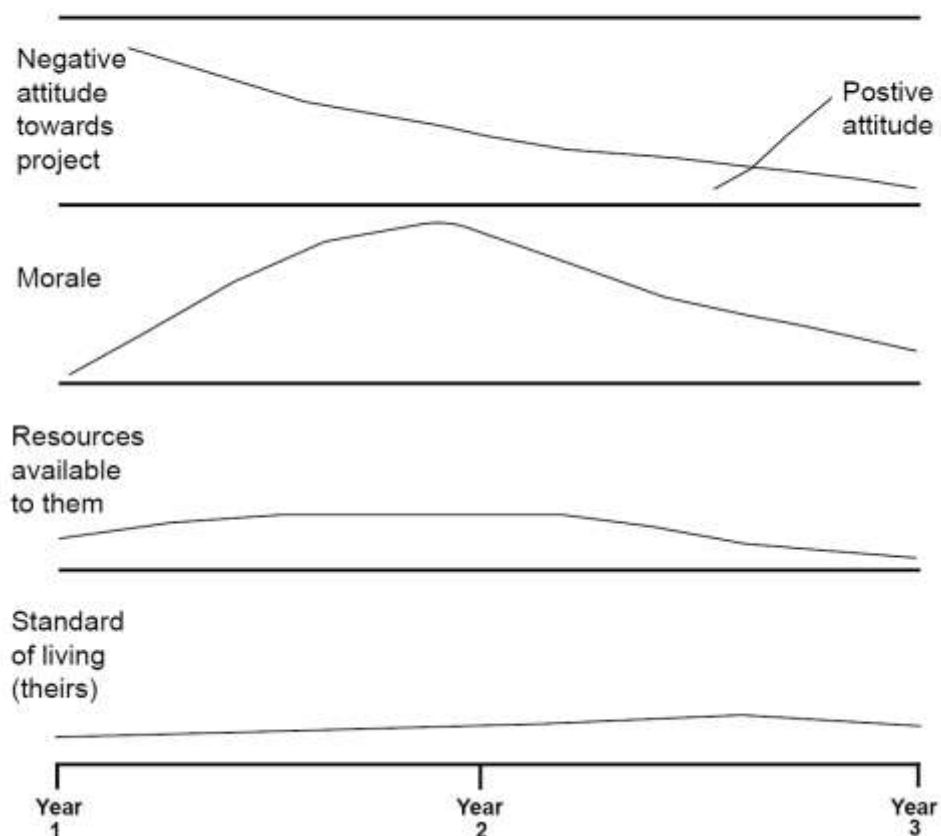
Those to whom management at the project should devolve.



**Figure 4: Government Divisional Officers**

**Stakeholders:**

Governments Divisional Officers



People opened up to us in surprising ways. One theme that emerged in our discussions was how expectations at the outset were stoked and maintained by a steady stream of white visitors. Two of us were white. But once our PRA methods got people focusing on the diagrams they were producing, they began telling us their stories; there seemed to be far less hesitation about being critical than would be the case in the usual kind of evaluation encounters involving white outsiders like ourselves. A time trends exercise with a community resource person revealed that his expectations of gaining something material from the program had been persistently thwarted, and remained a source of dissatisfaction. So when the program brought outside visitors and he had to spend the day without even being offered any food for the time he spent waiting around for them to arrive and talking to them, he came to resent it. The day before we visited him in his home, he told us, he had been called to meet with the team that were conducting the review:

We hear about X [the donor] giving grants and support all over the country, and they give the government allowances and people see that, it is done in front of them, yet there is nothing for us. Like yesterday when we were out from 10am to 4pm without taking anything to eat for that review team.

Why carry on, we asked him. At first, he told us, 'we thought the Whites were bringing the money'. He had waited for that money, and waited some more and then just said to himself, 'this is our country. Money or not money, let us get on'. But he was still waiting for some

acknowledgement from the program, and for the material assistance that he continues to expect for all that he sees the benefits of the program. The incident he narrated of being summoned to meet visitors is a development classic: in convening the obligatory performance of project success to visitors, scarce thought was given to the wellbeing of those putting on the performance, paradoxical indeed for a nutrition project to leave its participants hungry.

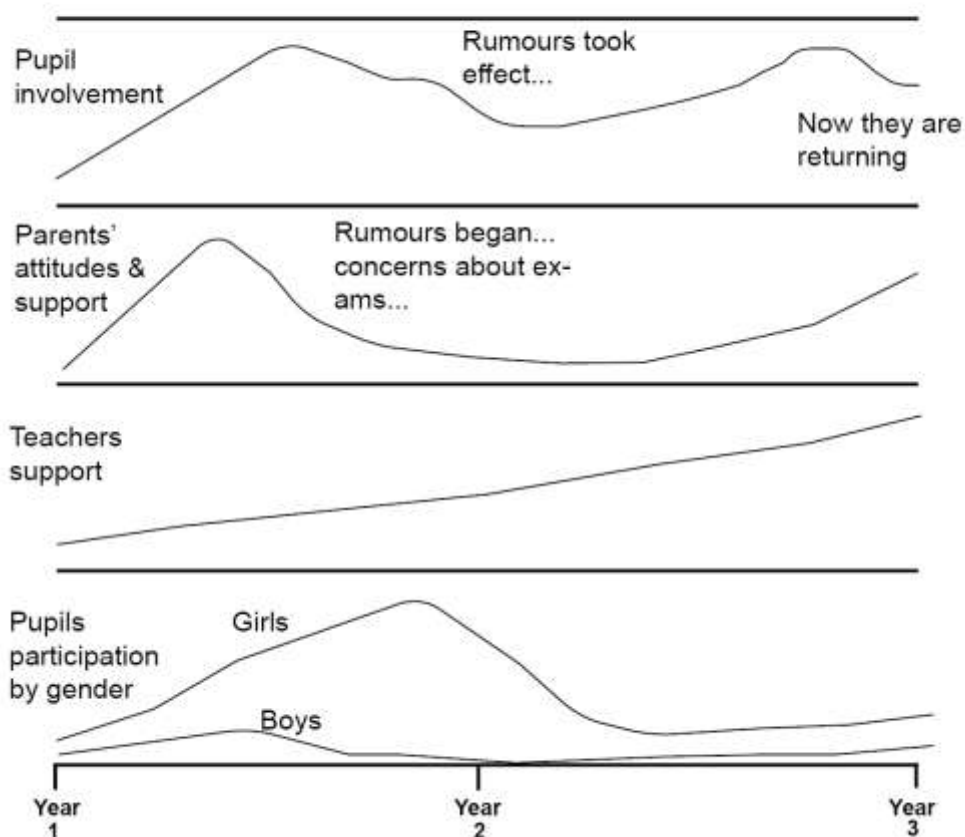
Children at one of the schools told us that boys complained and didn't want to take part in the program at the beginning because it involved having to sing to visitors. The boys didn't think this was something boys should be doing; it was 'girls work', they said, and in any case, they were shy about singing in public. They told us how children had stopped coming to the club run by the program because many parents were worried about the associations that rumour had made between the program and devil worship. When we talked to the teachers and program staff, another lesson learnt became evident: the program staff realized that it wasn't enough to seek children's participation through the schools and to regard the children as a conduit to changing behaviour at home. Their parents had to be brought on board. The decision to do so reversed the dip in involvement, and had other effects as parents began to notice changes in their children and attribute them to the program.

These changes also proved decisive in convincing teachers that the Child-to-Child clubs supported by the program were worth their support. Using diagrams and cards to have a conversation opened up a space for teachers to tell us about their initial scepticism, if not downright hostility, to the initiative; they feared it would distract the children from their studies and waste their time. One of the teachers confided that at the beginning, they thought they'd be given books, equipment and other things, and that the teachers would get money and become rich. They'd become accustomed to hearing about these kinds of things, and seeing handouts. The handouts didn't arrive. But the Child-to-Child activities began to have tangible, visible, effects on the children and the neighbouring community. Teachers told us how, slowly, they'd come around. They didn't expect the children to get so involved, they said, nor the kind of changes that they began to see, changes that exceeded anything they could have expected at the outset. Something that especially surprised them was seeing the children teaching themselves and others. 'The children have become quicker to learn for themselves', one teacher commented, so it became easier to teach them how. One of the parents remarked on how she'd seen the self-confidence of the children grow, as they learnt and taught each other: she talked of how she'd seen her daughter teaching her younger son, and both of them begin to help in the house to make some of the changes the program was teaching them about.

**Figure 5: Time-Trends: School Teachers**

**Stakeholders:**

School teachers (involves in child-to-child activities)



Many of our richest insights came from the application of the simple line diagram sequence described earlier. These methods offered a means of structuring an otherwise open-ended narrative biographical process in ways that generated comparable visual artefacts that could then become part of a meta-analysis of emergent themes. Putting the time trends diagrams generated in a given session side to side and looking for patterns, we could engage the stakeholder groups we worked with in reflecting on the complex causalities bound up with the changes that had taken place – negative, as well as positive, unexpected as well as expected. Our commitment to confidentiality, the sensitivities that emerged and our respect for the intimacy of confidences shared in these encounters all meant that we were unwilling to bring our informants together to compare versions. The task of analysis was, ultimately, ours. We worked through our materials with painstaking care, juxtaposing the diagrams and transcribed commentaries and looking inductively for patterns, then focusing on understanding the nuance of the situated versions told to us by different stakeholders. In this way, we formed a collective version of the 'stories of change' we were hearing.

These 'stories of change' provided insights into how different dimensions of the unfolding story intersected with each other. These simple line diagrams told their own stories; the interactions of impacts, including some of the changes that were taking place amongst different stakeholders in difficult-to-quantify things like trust, morale and enthusiasm, painted a complex picture in which the path of change was far from regular or linear. One of the arguably most impressive achievements of the program would not even have come into view

using conventional methods. Hearing stories that a neighbouring district had copied some of the practices introduced by the program, we requested a vehicle to see for ourselves. The program manager was hesitant: what did this have to do with the evaluation? *Everything*, we said. A bumpy ride later, we came upon a scene of garden plots of leafy green vegetables, dish racks and latrines, all ideas borrowed from the program area and diffused through people's own networks beyond the geographical boundaries of the program: evidence of impact that might otherwise have remained invisible.

## 5 Conclusion

Conventional impact assessment works on the basis of snapshots in time. Increasingly, it has come to involve reductionist metrics in which a small number of measurable indicators are taken to be adequate substitutes for the rich, diverse and complex range of ways in which change happens and is experienced. This kind of impact assessment is able to establish only that a change in the things that are measured took place. It is not able to offer much in the way of insight into what drives and sustains change. Howard White argues that:

Impact evaluations have become very important in recent years. If these studies are really to help improve development policies and programs, and so improve lives, then the studies need to go beyond just reporting a counterfactual-based measure of the impact on selected outcomes. Impact analysis needs to be embedded in a well-contextualised analysis that unpacks the causal chain. Such analysis helps understand why an intervention works or not, or why it only works for certain people, or in certain places or at certain times. (2011: 143)

Oakley et al. (2006:414) go further, describing how the use of process evaluation in conjunction with outcome data from an RCT, 'maximised our ability to interpret results according to empirical evidence.' They go on to make the case for enhancing RCTs by making process evaluation integral to them, suggesting that 'additional costs (such as collecting and analysing qualitative data) would probably be balanced by greater explanatory power and understanding of the generalisability of the intervention... A detailed process evaluation should be integral to the design of many randomised controlled trials' (2006: 415). They make the interesting argument that integrating process evaluation into RCTs helps persuade sceptics of the usefulness of experimental methods in the evaluation of complex development interventions.

What participatory process evaluation offers is not only a means through which to better understand 'impact by design'. It also, as I hope this article has shown, offers an opportunity to understand 'impact by interaction' and 'impact by emergence' (Bhola 2000). It can help to answer the question 'how could it be improved' (Jones et al. 2009), as well as 'will it work for us (elsewhere)' (Cartwright and Munro 2010), as well as some of the other questions Stern et al. (2012) highlight, for which experimental methods cannot essentially provide us with any credible answers. In the context described here, unanticipated effects were more important to the success of the program in the longer term than anyone imagined at the outset. Some of those effects came about through aspects of the process that had not been given any thought to in the design of the program. We came to learn, for example, how important rivalry and conflict were as a motivating force driving some of what could be claimed as success. Yet this would not – and could not – have been deliberately built into the program, nor would it have provided the basis for measurable indicators. And we came to see tangible signs of sustainable success in places that a conventional approach to evaluation would simply have placed out of view.

Participatory process evaluation can provide a powerful complement and corrective to conventional impact evaluation methodologies, reaching the parts that such methodologies leave inaccessible or out of view. By focusing on factors for success and on a multi-stakeholder narrative approach that is capable of explaining how and why change happens, and by exploring 'impact trajectories' (Woolcock 2009) in a way that is sensitive to positionality and dimensions of impact that are otherwise difficult to access, participatory process evaluation offers a more rigorous and thoroughgoing methodology for understanding the dynamics of change than impact evaluation that relies on quantification alone. By involving those who are engaged at different level in program implementation in reflecting on their own practice and their learning journeys, this approach can provide a valuable space for reflexive learning. It offers considerable scope for building practitioners' capacities in reflective practice. As such, it could be incorporated into developing communities of practice within the context of ongoing program monitoring, as well as through periodic evaluations.

## References

- Bhola, H. (2000) 'A Discourse on Impact Evaluation: A Model and its Application to a Literacy Intervention in Ghana', *Evaluation* 6.2: 161–78
- Bleek, Wolf (1987) 'Lying Informants: A Fieldwork Experience from Ghana', *Population and Development Review* 13.2: 314–22
- Cartwright, N. and Munro, E. (2010) 'The Limitations of Randomized Controlled Trials in Predicting Effectiveness', *Journal of Evaluation in Clinical Practice* 16: 260–6
- Carvalho, S. and White, H. (1997) *Combining the Quantitative and Qualitative Approaches to Poverty Measurement and Analysis. The Practice and the Potential*, World Bank Technical Paper 336, Washington DC: World Bank
- Chambers, Robert (1997) *Whose Reality Counts: Putting the First Last*, London: Intermediate Technology Publications
- Chambers, Robert and Mayoux, Linda (2005) 'Reversing the Paradigm: Quantification, Participatory Methods and Pro-poor Impact Assessment', *Journal of International Development* 17: 271–98
- European Evaluation Society (2007) *The Importance of a Methodologically Diverse Approach to Impact Assessment*, [http://www.europeanevaluation.org/images/file/Member\\_section/EES%20Statement%20on%20Methodological%20Diversity/EES%20Statement.pdf](http://www.europeanevaluation.org/images/file/Member_section/EES%20Statement%20on%20Methodological%20Diversity/EES%20Statement.pdf) (accessed 15 April 2013)
- Eyben, Rosalind (2013) *Uncovering the Politics of 'Evidence' and 'Results'. A Framing Paper for Development Practitioners*, Brighton: Big Push Forward
- Gaventa, John (1993) 'The Powerful, the Powerless and the Experts: Knowledge Struggles in a Information Age' in Peter Park, Budd Hall and Ted Jackson (eds), *Participatory Research in North America*, Amherst MA: Bergin and Hadley
- Guijt, Irene (2008a) 'Seeking Surprise: Rethinking Monitoring for Collective Learning in Rural Resource Management', PhD Thesis, Wageningen University, Holland
- Guijt, Irene (2008b) *Critical Readings on Assessing and Learning for Social Change: A Review*, IDS Development Bibliography 21, Brighton: IDS

- Havemann, K.; Pridmore, P.; Tomkins, A. and Garn, K.D. (2012) 'What Works and Why? Evaluation of a Community Nutrition Programme in Kenya', *Public Health Nutrition* 19: 1–8
- Jones, H. (2009) 'The "Gold Standard" is not a Silver Bullet', *ODI Opinion Piece* 127, London: ODI
- Jones, N.; Jones, H.; Steer, L. and Datta, A. (2009) *Improving Impact Evaluation Production and Use*, Occasional Paper 300, London: Overseas Development Institute (ODI), [www.odi.org.uk/resources/docs/4158.pdf](http://www.odi.org.uk/resources/docs/4158.pdf)
- Mackay, R. and Horton, D. (2003) 'Expanding the Use of Impact Assessment and Evaluation in Agricultural Research and Development', *Agricultural Systems* 78.2: 143–65
- Oakley, A.; Strange, V.; Bonell, C.; Allen, E. and Stephenson, J. (2006) 'Process Evaluation in Randomised Controlled Trials of Complex Interventions', *British Medical Journal* February 18; 332(7538): 413–6
- Ravallion, M. (2009) 'Should the Randomistas Rule?', *The Economists' Voice* 6.2: 1–5
- Roche, C. (2000) 'Impact Assessment: Seeing the Wood and the Trees', *Development in Practice* 10.3&4: 543–55
- Stern, E.; Stame, N.; Mayne, J.; Forss, K.; Davies, R. and Befani, B. (2012) *Broadening the Range of Designs and Methods for Impact Evaluations*, Working Paper 38, Department for International Development (DFID)
- Tandon, Rajesh (1996) 'The Historical Roots and Contemporary Tendencies in Participatory Research: Implications for Health Care', in Korrie de Koning and Marion Martin (eds), *Participatory Research in Health: Issues and Experiences*, London: Zed Books
- White, Howard (2011) 'Achieving High-quality Impact Evaluation Design through Mixed Methods: The Case of Infrastructure', *Journal of Development Effectiveness* 3.1, March: 131–44
- White, Howard (2008) 'Of Probits and Participation: The Use of Mixed Methods in Quantitative Impact Evaluation', *IDS Bulletin* 39.1: 98–109
- Woolcock, M. (2009) *Towards a Plurality of Methods in Project Evaluation: A Contextualized Approach to Understanding Impact Trajectories and Efficacy*, Working Paper 73, University of Manchester: Brooks World Poverty Institute